UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 09/991,127 | 11/14/2001 | Ethan George Russell | 55994.0120 | 5209 |

57600 7590 04/29/2008

HOLLAND & HART LLP
60 E. South Temple, Suite 2000
P.O. Box 11583
Salt Lake City, UT 84110

| EXAMINER | |
|---|---|
| STRANGE, AARON N | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2153 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 04/29/2008 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 09/991,127 | RUSSELL ET AL. | |
| | Examiner | Art Unit | |
| | AARON STRANGE | 2153 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *05 February 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-8,10,12,21-29,31-38,40-50,52-56 and 62-68* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-8,10,12,21-29,31-38,40-50,52-56 and 62-68* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

## DETAILED ACTION

### *Response to Arguments*

1. Applicant's amendments to claims 21-34 and 62-68 are sufficient to overcome the rejection of those claims set forth under 35 U.S.C. §112, second paragraph, and that rejection has been withdrawn.

2. Applicant's arguments filed 2/5/08 have been fully considered but they are not persuasive.

3. With regard to Applicant's assertion that the combined teachings of Bryant, Bland and Wang fail to disclose, teach or suggest "determining whether distributed application data (or a Web page) have been cached at a second site (or by a client device) prior to determining a performance metric" (Remarks 22), the Examiner respectfully disagrees.

As discussed below, Bryant teaches determining the response time of a web server and that the response time is the time period between a client making a request and receiving a response (col. 6, ll. 49-52). Bryant discloses that this information is used by Web site operators to evaluate the quality of service (col. 2, ll. 15-18).

Wang teaches the use of caching to improve the quality of service for requesting clients, since caching can reduce the access latency experienced by clients (§3 & 4.1.1). The system determines whether a request may be satisfied by a local cache prior to forwarding the request to another cache or the source server (§4.1.1). These

types of caching systems were old and well known in the art at the time the invention was made.

One of ordinary skill in the art, aware of the widespread use of caching systems and their purpose of reducing the access latency for clients, would have recognized that clients requesting cached copies of Bryant's web pages would have measured a smaller response time than they would have measured if the page had not been cached. This measurement would have been an inaccurate representation of the actual response time experienced by clients without cached copies of the distributed application data.

Upon recognizing this problem, the combined teachings of Bryant and Wang would have suggested to one of ordinary skill in the art that determining whether the requested data is already cached at the client would be beneficial, since the response time measured for cached data is an inaccurate representation of the response time that would have been experienced if the data had not been cached. Since the response time information is intended for use by Web site operators to evaluate quality of service (Bryant, col. 2, ll. 15-18), it would be beneficial to know whether low response times are a function of the Web site's performance or the requested data being cached.

### *Claim Rejections - 35 USC § 103*

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claims 1-8, 10, 12, 21-29, 31-38, 40-50, 52-56 and 62-68 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bryant et al. (US 6,078,956) in view of Bland et al. (US 5,732,218) further in view of Jia Wang ("A Survey of Web Caching Schemes for the Internet").

6. With regard to claim 1, Bryant discloses a method for determining one or more performance metrics for a distributed application in which distributed application data are transferred from a first site (server) to a second site (client) over a network, comprising the steps of:

(a) enabling a user to transmit a request for the distributed application data desired by the user, said request being transmitted from the second site to the first site over the network (user clicks link to submit request)(at least Col 7, Line 61 to Col 8, Line 13);

(b) in response to the request, transmitting the distributed application data from the first site to the second site over the network, if the distributed application data are not already accessible at the second site (link is followed) (at least Col 8, Lines 16-17);

(c) including machine instructions that define a performance monitoring function with the distributed application data that were requested and transmitted over the network to the second site, the machine instructions and the distributed application data being a single data file (code is submitted as part of the page)(at least Col 8, Line 66 to Col 9, Line 15); and

(d) executing the machine instructions at the second site(received script is run), to implement the performance monitoring function used to determine the one or more performance metrics (time required to service request) (at least Col 9, Lines 5-18) for the distributed application without using the performance monitoring function to request any distributed application data from any site (no requests are made in code), at least one of the one or more performance metrics being determined in connection with timing of events occurring during the transmission of the distributed application data to the second site (the time to download the page containing the distributed application data) (at least Col 9, Lines 5-18).

Bryant further discloses how correlated performance metrics may be determined by combining the one or more performance metrics determined at the second site with a performance metric determined at the first site (col. 5, ll. 7-26), but fails to specifically disclose doing so. Bryant also fails to disclose determining whether the distributed application data (web page) has already been cached at the second site (client), before determining a performance metric.

Bland discloses a similar system for monitoring interactions between a first site and a second site (col. 2, ll. 1-16). Bland teaches determining a first performance metric at a first site (server collects performance metrics such as server processing time)(col. 3, ll. 17-22 & 48-52), determining a second performance metric at a second site (client collects metrics such as delays between requests and responses)(col. 3, ll. 23-29 and col. 4, ll. 11-12), and combining the metrics to determine a correlated performance metric (subtraction of server processing time from total delay results in the network

delay time)(col. 4, ll. 15-20). This would have been an advantageous addition to the system disclosed by Bryant since it would have allowed more detailed information about the network to be gathered, and assisted in identifying the causes of delays experienced by users.

Wang discloses that browser caches and proxy caches are well-known in the art for maintaining local copies of web documents (Section 4.1.1). Browser caches and proxy caches provide reduced latency for accessing web documents (Section 3). The closer the document is to the requesting client, the faster it will be able to retrieve it. It would have been advantageous to determine if the distributed application data being requested by the client had been cached prior to determining a performance metric. If the data hds been cached, the latency will be significantly lower than it would have been for uncached data, making the collected data an inaccurate representation of the Web site's performance and network latency.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to collect performance metrics at both the client and the server, and combine the metrics to determine correlated performance metrics such as network delay time that cannot be determined by monitoring at a single site as well as determine if the requested data has been cached prior to determining a performance metric since cached data will be retrieved much more quickly and the results will not be an accurate indicator of Web site performance and actual network latency.

7. With regard to claim 2, Bryant further discloses that the performance monitoring function at the second site is initiated after the distributed application data are accessed at the second site (performance monitoring function is received appended to the distributed application data)(at least Col 9, Lines 5-18).

8. With regard to claim 3, Bryant further discloses the step of collecting the one or more performance metrics for the distributed application over the network (at least Col 9, Lines 19-28).

9. With regard to claim 4, Bryant further discloses applying a probabilistic sampling parameter to determine whether performance metrics are collected from each of a plurality of sites (at least Col 6, Lines 41-46).

10. With regard to claim 5, Bland further discloses collecting data about client sessions wherein data about request delays is collected for an entire session prior to sending it to the server (Col 3, Lines 19-23).

11. With regard to claim 6, Bryant further discloses that the probabilistic sampling parameter is applied on a per-request basis (individual request times are sampled) (at least Col 6, Lines 31-46).

12. With regard to claim 7, Bryant further discloses that the performance monitoring function at the second site determines at least: (a) a fetch latency, corresponding to a time period required to fetch the distributed application data (at least Col 9, Lines 5-18).

13. With regard to claim 8, Bland further discloses determining whether to collect a performance metric as a based on a specific kind of performance metric that was determined (clients only collect data that is pertinent to a server in response to a request from that server) (Col 4, Line 64 to Col 5, Line 16).

14. With regard to claim 10, Bryant further discloses the distributed application data have a markup language format (web pages) (at least Col 8, Line 61 to Col 9, Line 13).

15. With regard to claim 12, Bryant further discloses that said one or more performance metrics is determined without any apparent effect on the access of the distributed application data at the second site (metrics are determined after page is retrieved) (at least Col 9, Line 15-28).

16. With regard to claim 21, Bryant discloses a method for determining and collecting at least one performance metric related to access of a Web page by a browser program on a client device, including at least one of a compound performance metric and a correlated performance for a network, comprising the steps of:

(a) enabling a user to request transfer of the Web page from a server device to the client device over a network (user clicks link to submit request)(at least Col 7, Line 61 to Col 8, Line 13);

(d) including machine instructions with the Web page so that the Web page and machine instructions are transferred to the client device as one data file (code is submitted as part of the page)(at least Col 8, Line 66 to Col 9, Line 15);

(e) when the Web page is loaded by the client device for rendering by the browser program, causing the client device to execute the machine instructions that define how to carry out the browser monitoring function, said browser monitoring function being implemented without requiring any affirmative action by a user of the client device (metrics are automatically determined after page is retrieved) (at least Col 9, Line 15-28);

(f) determining said at least one performance metric on the client device with the browser monitoring function without using the browser monitoring function to request any Web page from any site, at least one performance metric being determined in connection with timing of events occurring during transmission of the distributed application data to the client device (at least Col 9, Line 15-28); and

(e) determining a server performance metric; and

(f) combining the server performance metric with said at least one performance metric to determine the correlated performance metric (Internet Delay is combination of SERVER_PROCESSING_DELAY and response time)(at least Col 5, Lines 7-26).

Additionally, it should be noted that step (e) is an optional step, only required when a correlated performance metric is to be determined.

Bryant further discloses how correlated performance metrics may be determined by combining the one or more performance metrics determined at the second site with a performance metric determined at the first site (col. 5, ll. 7-26), but fails to specifically disclose doing so. Bryant also fails to disclose determining whether the distributed application data (web page) has already been cached at the second site (client), before determining a performance metric.

Bland discloses a similar system for monitoring interactions between a first site and a second site (col. 2, ll. 1-16). Bland teaches determining a first performance metric at a first site (server collects performance metrics such as server processing time)(col. 3, ll. 17-22 & 48-52), determining a second performance metric at a second site (client collects metrics such as delays between requests and responses)(col. 3, ll. 23-29 and col. 4, ll. 11-12), and combining the metrics to determine a correlated performance metric (subtraction of server processing time from total delay results in the network delay time)(col. 4, ll. 15-20). This would have been an advantageous addition to the system disclosed by Bryant since it would have allowed more detailed information about the network to be gathered, and assisted in identifying the causes of delays experienced by users.

Wang discloses that browser caches and proxy caches are well-known in the art for maintaining local copies of web documents (Section 4.1.1). Browser caches and proxy caches provide reduced latency for accessing web documents (Section 3). The

closer the document is to the requesting client, the faster it will be able to retrieve it. It would have been advantageous to determine if the distributed application data being requested by the client had been cached prior to determining a performance metric. If the data hds been cached, the latency will be significantly lower than it would have been for uncached data, making the collected data an inaccurate representation of the Web site's performance and network latency.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to collect performance metrics at both the client and the server, and combine the metrics to determine correlated performance metrics such as network delay time that cannot be determined by monitoring at a single site as well as determine if the requested data has been cached prior to determining a performance metric since cached data will be retrieved much more quickly and the results will not be an accurate indicator of Web site performance and actual network latency.

17. With regard to claim 28, Bland further discloses that determining said at least one performance metric is done without the client device providing any indication to the user of the client device that said at least one performance metric is being determined (parameters may be collected by the client automatically for all data or in response to a request from a server) (Col 4, Lines 60-67). Bland also teaches that notifying and requesting permission before collecting data is optional (Col 5, Lines 11-14).

18. With regard to claim 31, Bryant further discloses that said at least one performance metric comprises a performance metric for each image included in the web page (response time of all gif's received)(at least Col 7, Lines 35-46 and Col 2, Lines 55-60).

19. With regard to claim 32, Bryant further discloses

including a monitor cookie with the web page that is transferred to the client device from the server device which indicates that the Web page is a monitored document;

detecting the monitor cookie when the Web page is transferred to the client device; and

causing the browser function to determine that said at least one performance metric is to be determined for the Web page in response to the monitor cookie being detected (at least Col 8, Lines 1-21).

20. With regard to claim 54, Bland further discloses collecting performance metrics for the server related to the transfer of Web pages (Col 3, Line 40 to Col 4, Line 59) and transmitting the metrics to a remote data center site for processing (central server that has management system) (Col 3, Lines 17-22).

21. With regard to claim 55, Bryant further discloses combining a performance metric determined by the browser monitoring function executed by the processing device with

the server performance metric determined by the server computing function to determine the correlated performance metric (at least Col 4, Line 65 to Col 5, Line 21 and Col 8, Lines 20-25).

22. With regard to claim 56, Bryant further discloses a caching proxy (proxy server 30) (Par 119, Lines 6-11) disposed between the server computing device and the processing device (Fig 1, 30), but fails to disclose said caching proxy executing a caching proxy monitoring function that determined at least one performance metric related to a performance of the caching proxy.

Bland teaches a method of collecting performance metrics for a server related to Transfer of Web page requests to a client. Bland discloses that several types of metrics are collected at the server (Col 3, Line 41 to Col 4, Line 59). For example, the delay between a client request and a server response is measured to determine the load on the server (Col 3, Lines 47-51). This would have been an advantageous addition to the system disclosed by Bryant since the proxy server can have a significant effect on the overall latency of client requests, and determining information about its performance is crucial to finding bottlenecks in the network.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to executing a monitoring function on the proxy server to determine at least one performance metric related to the performance of the caching proxy, since the proxy server has a significant effect on latency of client requests.

23. The remaining claims are rejected under the same rationale as the claims rejected above, since they recite substantially identical subject matter. Any differences between the claims do not result in patentably distinct claims and all of the limitations are taught by the above cited art. The table below lists the correlation between claims not individually rejected and the claims rejected above.

| | |
|---|---|
| 22, 23, 33, 34, 36, 43, 44, 63 and 68 | 3 |
| 24, 24 and 64 | 4 |
| 25 and 46 | 5 |
| 26 and 47 | 6 |
| 29, 38, 50 and 66 | 7 |
| 27, 48 and 65 | 8 |
| 62, 35 and 42 | 21 |
| 37 and 49 | 28 |
| 40, 52 and 67 | 31 |
| 41 and 53 | 32 |

## *Conclusion*

24. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to AARON STRANGE whose telephone number is (571)272-3959. The examiner can normally be reached on M-F 8:30-5:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Glen Burgess can be reached on 571-272-3949. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Glenton B. Burgess/

Supervisory Patent Examiner, Art Unit 2153

/A. S./
Examiner, Art Unit 2153